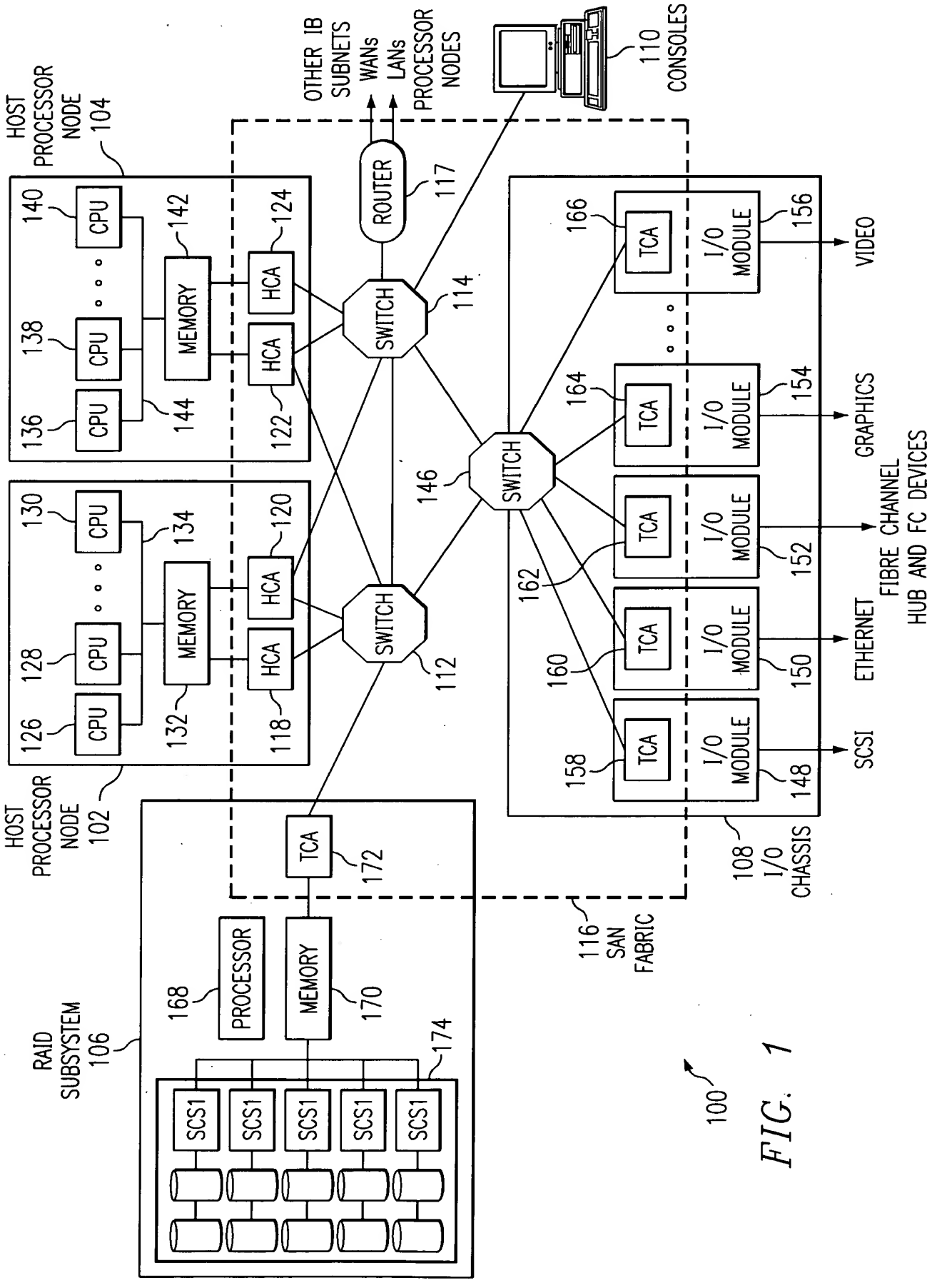

HOST PROCESSOR NODE 104
HOST PROCESSOR NODE 102
CPU
MEMORY
HCA
SWITCH
ROUTER
OTHER IB SUBNETS
WANs
LANs
PROCESSOR NODES
110 CONSOLES
TCA
I/O MODULE
VIDEO
GRAPHICS
ETHERNET
FIBRE CHANNEL HUB AND FC DEVICES
SCSI
108 I/O CHASSIS
116 SAN FABRIC
RAID SUBSYSTEM 106
PROCESSOR
SCSI
100

FIG. 1

FIG. 2

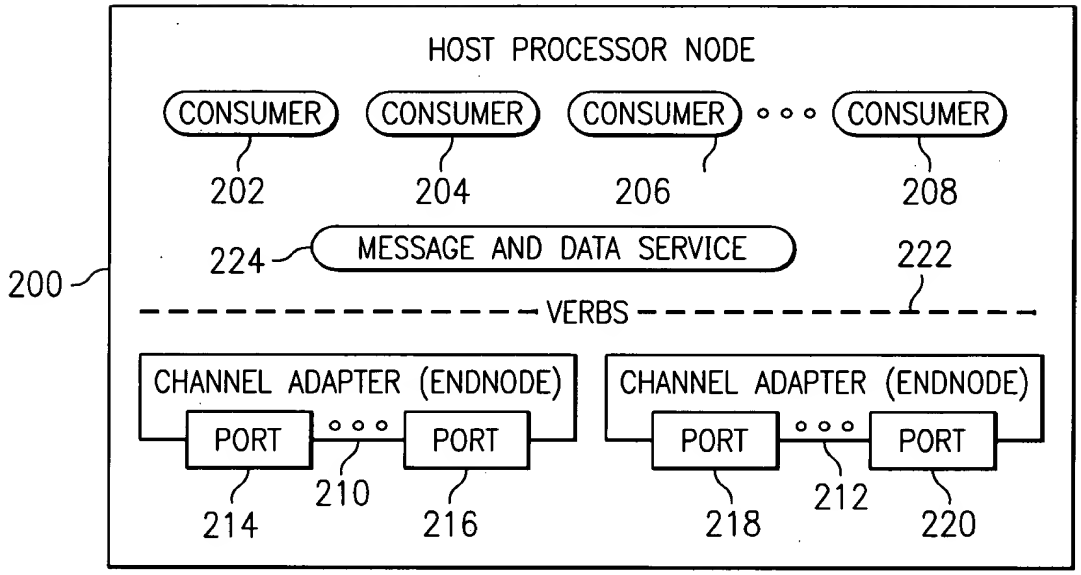

HOST PROCESSOR NODE
CONSUMER
CONSUMER
CONSUMER
CONSUMER
202
204
206
208
224
MESSAGE AND DATA SERVICE
222
200
VERBS
CHANNEL ADAPTER (ENDNODE)
PORT
PORT
210
214
216
CHANNEL ADAPTER (ENDNODE)
PORT
PORT
212
218
220

FIG. 3A

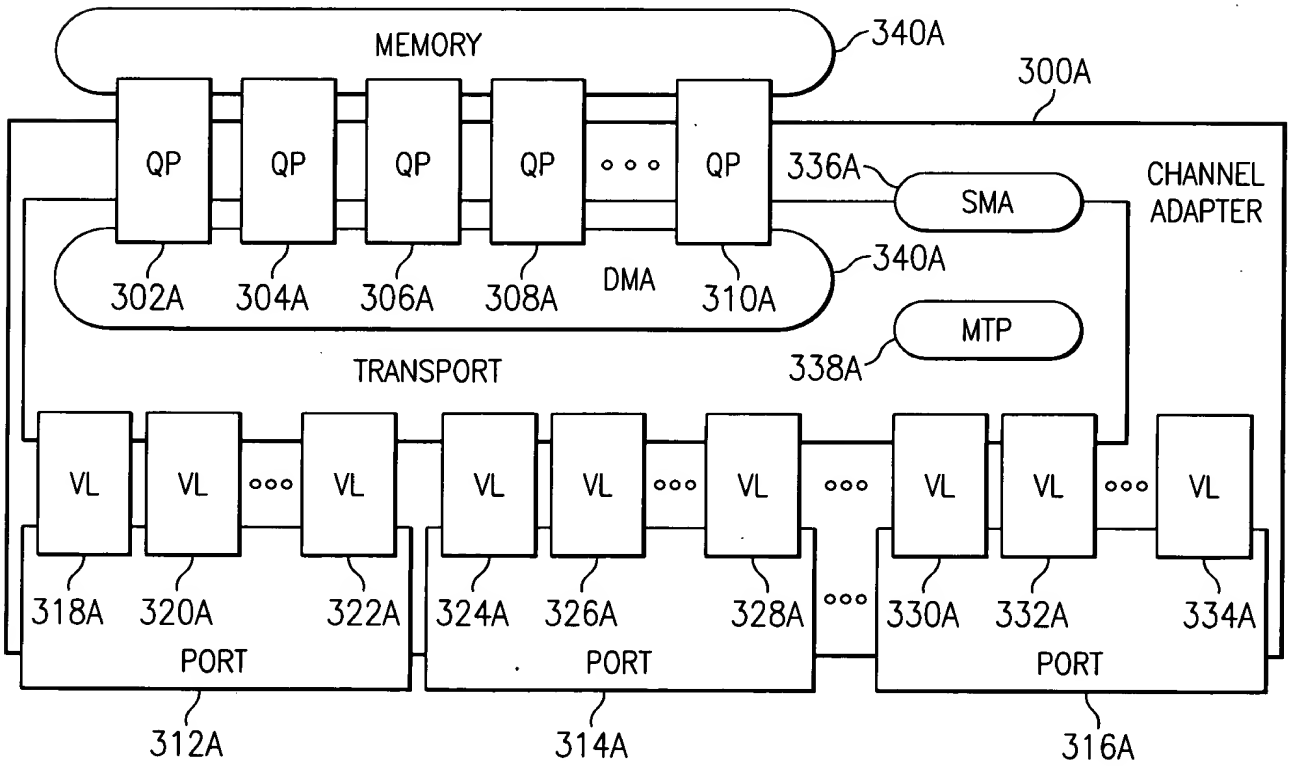

MEMORY
340A
300A
QP
QP
QP
QP
QP
336A
SMA
CHANNEL ADAPTER
340A
302A
304A
306A
308A
DMA
310A
MTP
338A
TRANSPORT
VL
VL
VL
VL
VL
VL
VL
VL
VL
318A
320A
322A
324A
326A
328A
330A
332A
334A
PORT
PORT
PORT
312A
314A
316A

FIG. 3B

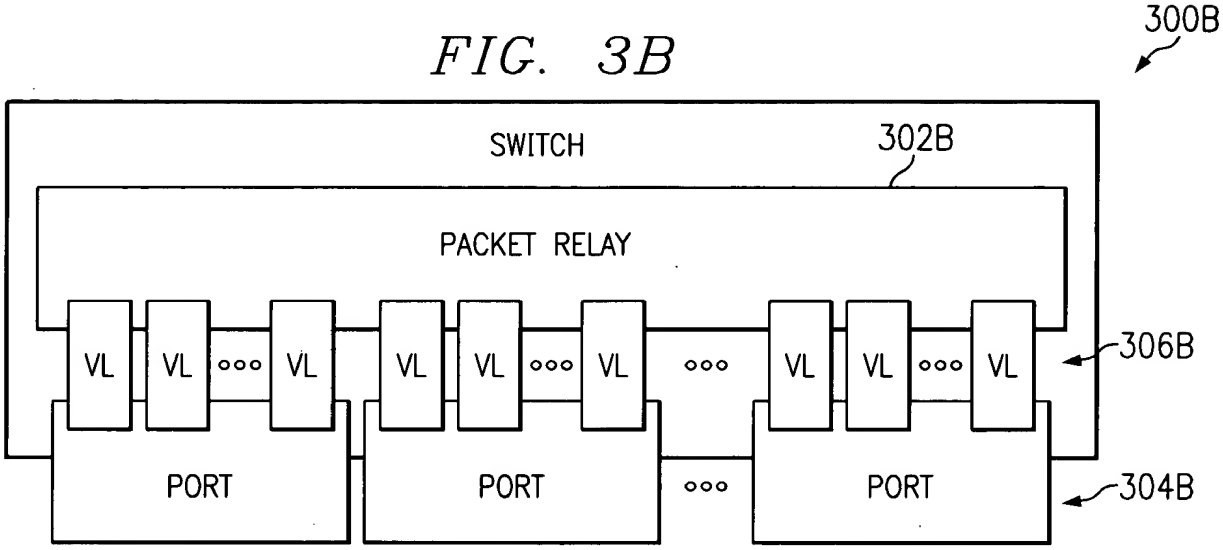
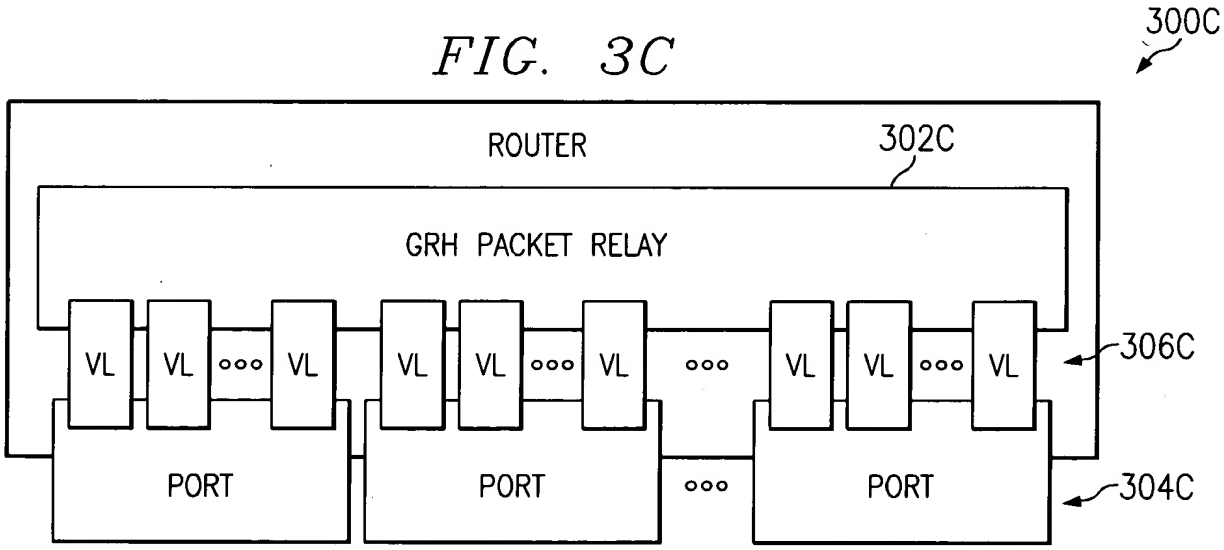

300B
SWITCH
302B
PACKET RELAY
VL
VL
VL
VL
VL
VL
VL
VL
VL
VL
306B
PORT
PORT
PORT
304B

FIG. 3C


300C
ROUTER
302C
GRH PACKET RELAY
VL
VL
VL
VL
VL
VL
VL
VL
VL
VL
306C
PORT
PORT
PORT
304C

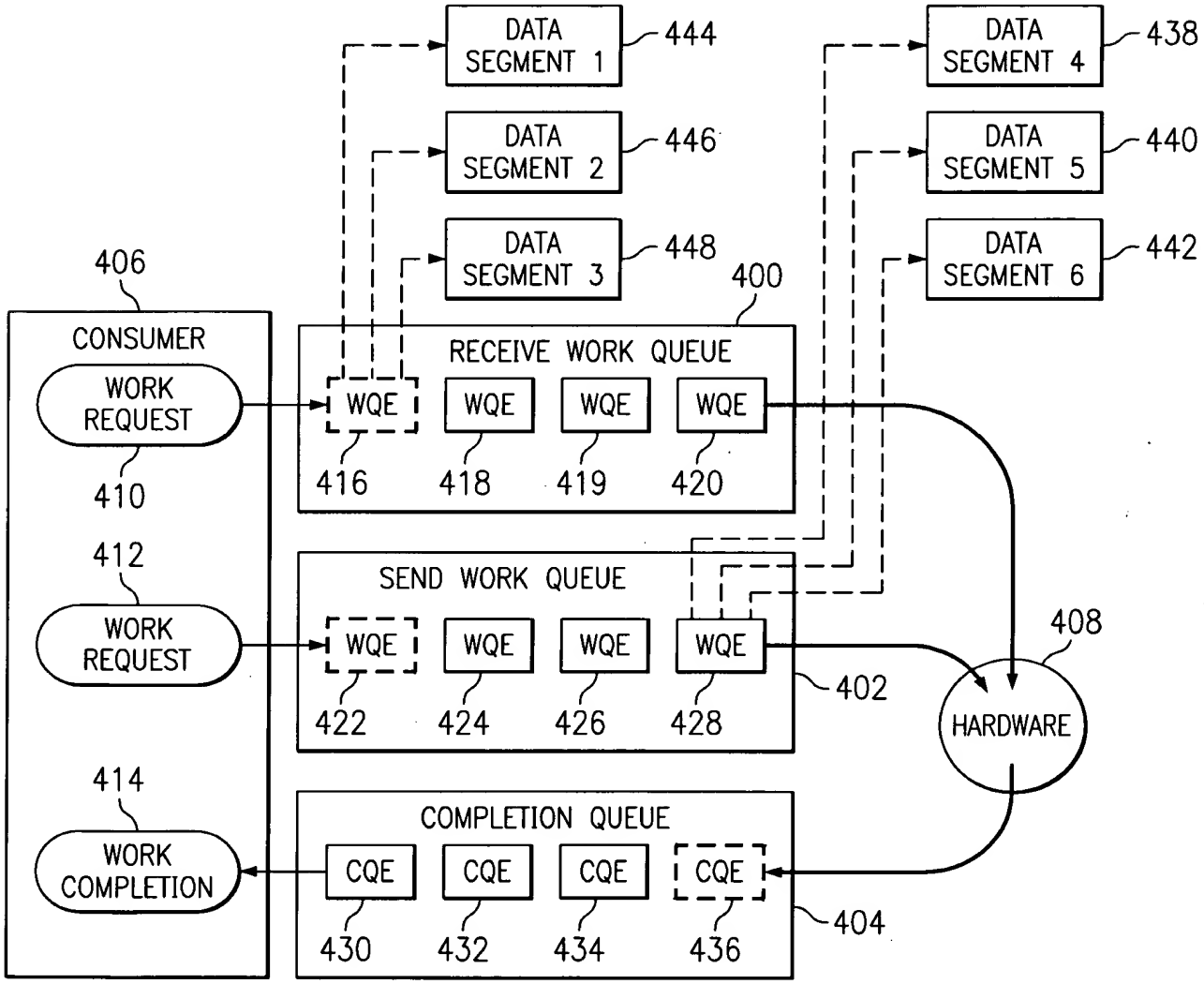

DATA SEGMENT 1
444
DATA SEGMENT 2
446
DATA SEGMENT 3
448
DATA SEGMENT 4
438
DATA SEGMENT 5
440
DATA SEGMENT 6
442
406
CONSUMER
WORK REQUEST
410
412
WORK REQUEST
414
WORK COMPLETION
400
RECEIVE WORK QUEUE
WQE
WQE
WQE
WQE
416
418
419
420
SEND WORK QUEUE
WQE
WQE
WQE
WQE
422
424
426
428
402
408
HARDWARE
COMPLETION QUEUE
CQE
CQE
CQE
CQE
430
432
434
436
404

FIG. 4

FIG. 5

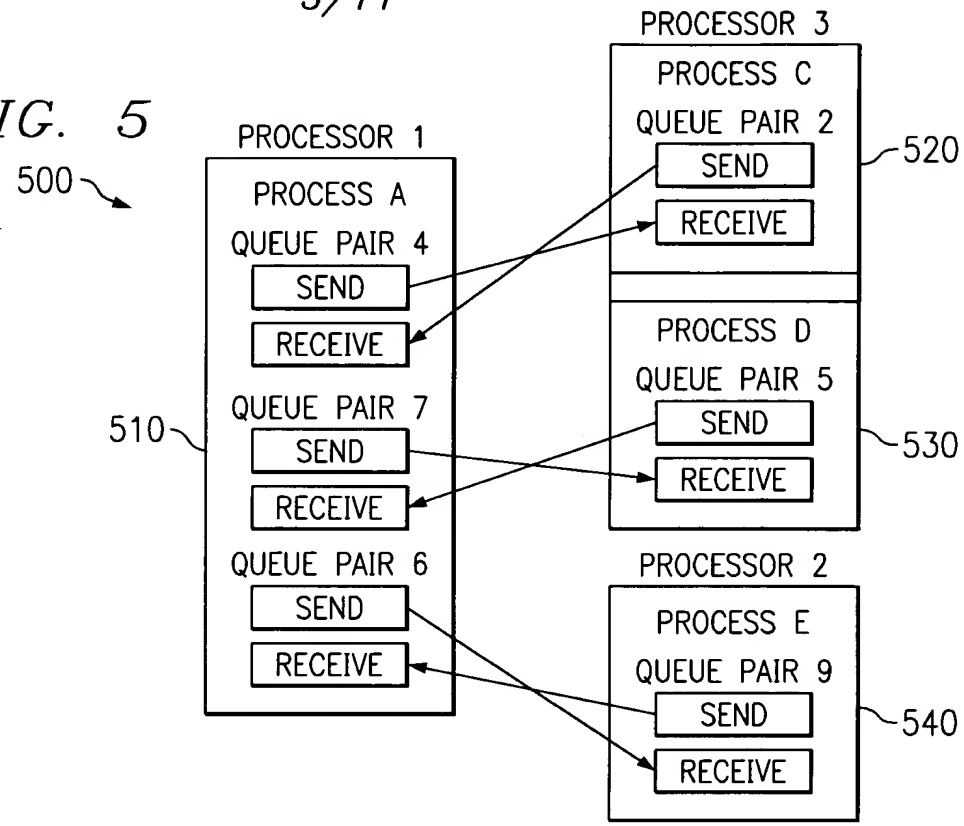

500
PROCESSOR 1
PROCESS A
QUEUE PAIR 4
SEND
RECEIVE
QUEUE PAIR 7
SEND
RECEIVE
QUEUE PAIR 6
SEND
RECEIVE
510
PROCESSOR 3
PROCESS C
QUEUE PAIR 2
SEND
RECEIVE
520
PROCESS D
QUEUE PAIR 5
SEND
RECEIVE
530
PROCESSOR 2
PROCESS E
QUEUE PAIR 9
SEND
RECEIVE
540

FIG. 6

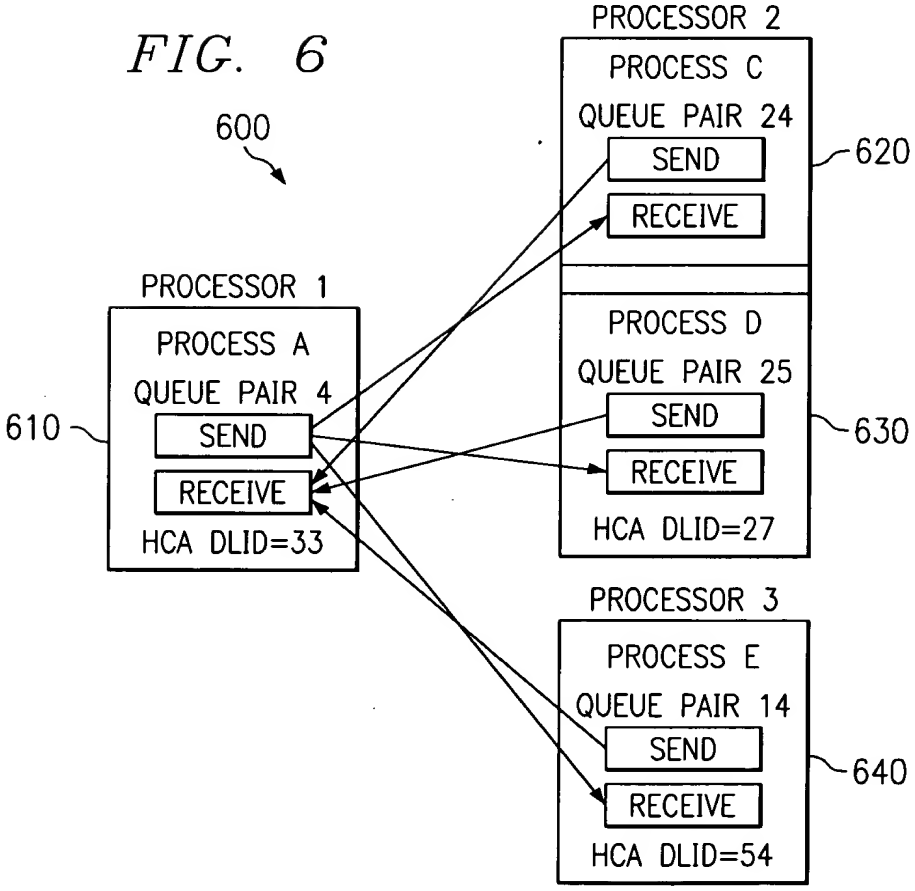

600
PROCESSOR 1
PROCESS A
QUEUE PAIR 4
SEND
RECEIVE
HCA DLID=33
610
PROCESSOR 2
PROCESS C
QUEUE PAIR 24
SEND
RECEIVE
620
PROCESS D
QUEUE PAIR 25
SEND
RECEIVE
HCA DLID=27
630
PROCESSOR 3
PROCESS E
QUEUE PAIR 14
SEND
RECEIVE
HCA DLID=54
640

FIG. 7

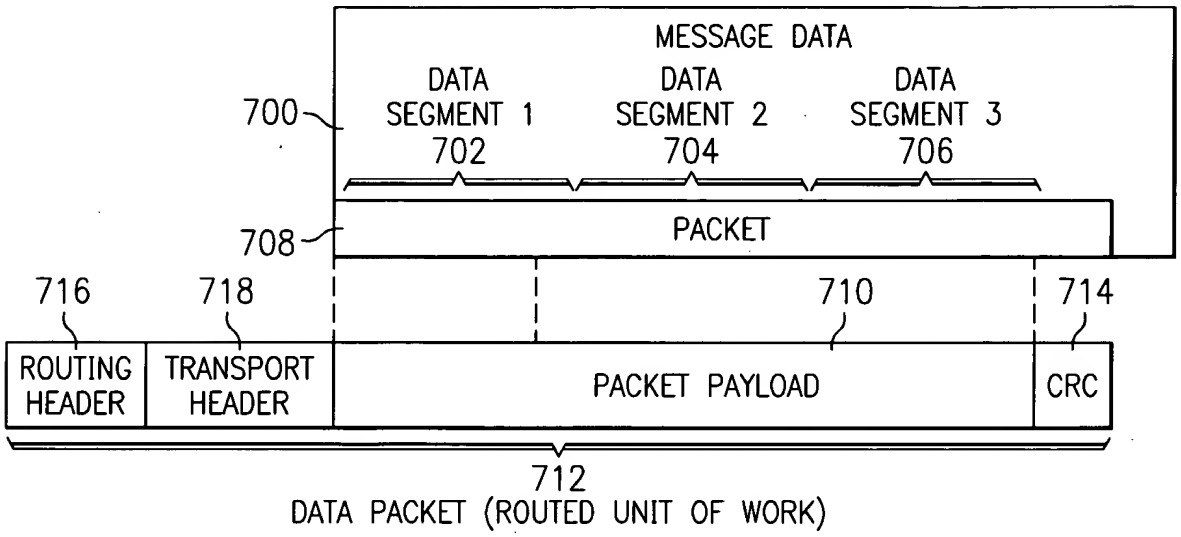

MESSAGE DATA
DATA SEGMENT 1
DATA SEGMENT 2
DATA SEGMENT 3
700
702
704
706
PACKET
708
716
718
710
714
ROUTING HEADER
TRANSPORT HEADER
PACKET PAYLOAD
CRC
712
DATA PACKET (ROUTED UNIT OF WORK)

FIG. 8

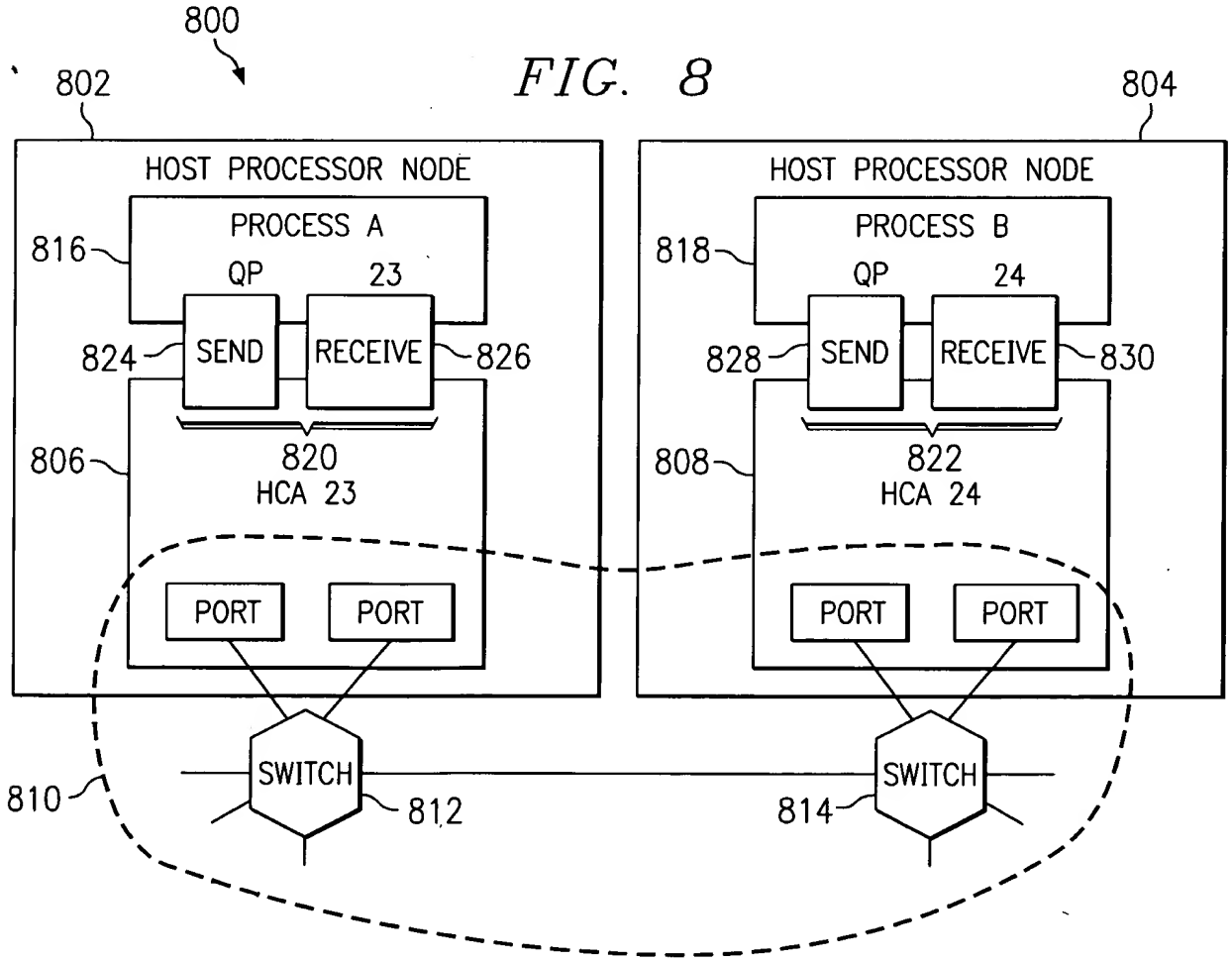
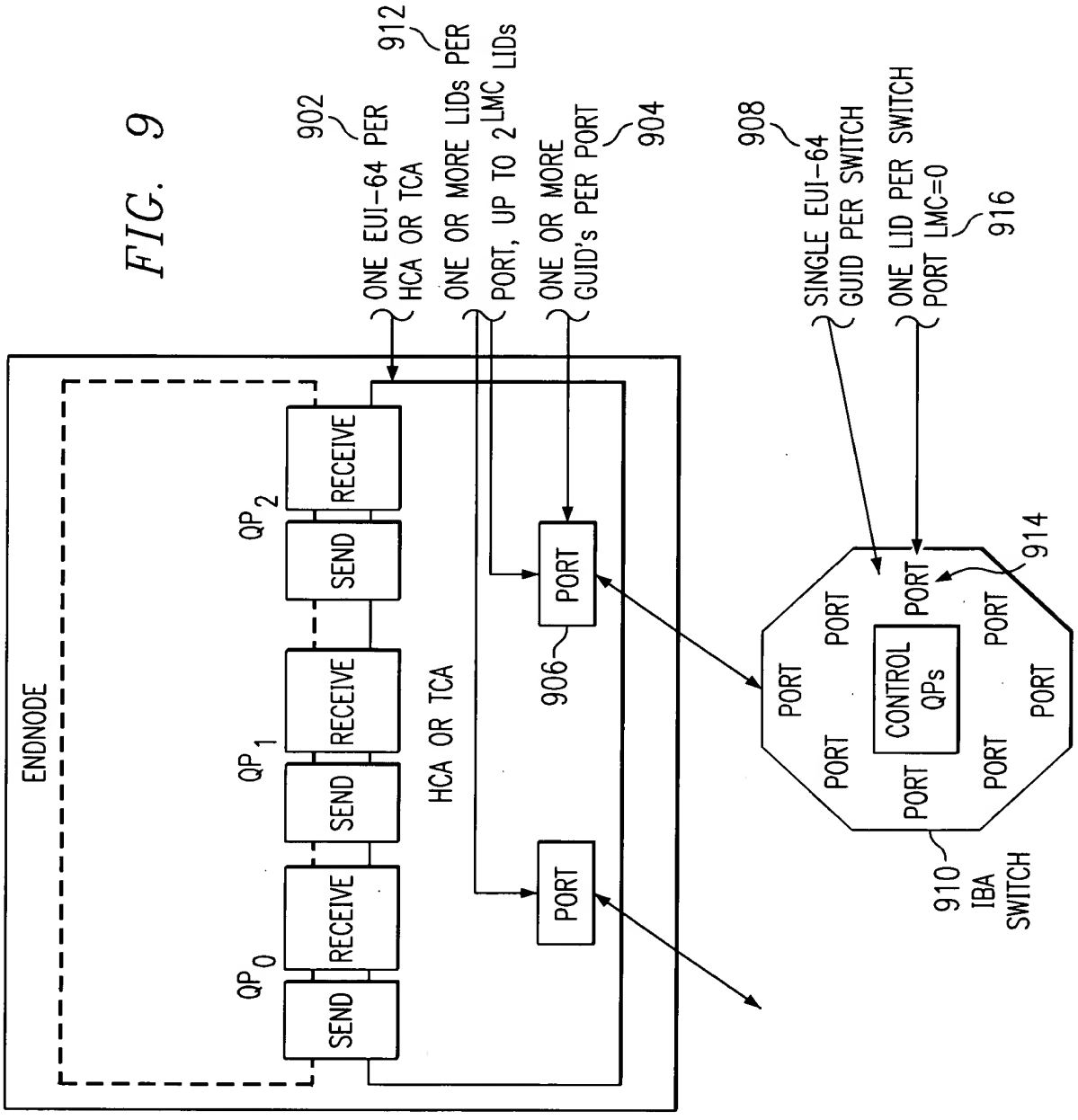

800
802
804
HOST PROCESSOR NODE
HOST PROCESSOR NODE
PROCESS A
PROCESS B
816
818
QP 23
QP 24
824
SEND
RECEIVE
826
828
SEND
RECEIVE
830
820
HCA 23
822
HCA 24
806
808
PORT
PORT
PORT
PORT
SWITCH
SWITCH
810
812
814

FIG. 9


902
ONE EUI-64 PER HCA OR TCA
912
ONE OR MORE LIDs PER PORT, UP TO 2LMC LIDs
904
ONE OR MORE GUID's PER PORT
908
SINGLE EUI-64 GUID PER SWITCH
ONE LID PER SWITCH PORT LMC=0
916
ENDNODE
HCA OR TCA
QP2
QP1
QP0
RECEIVE
SEND
PORT
906
914
CONTROL QPs
910
IBA SWITCH

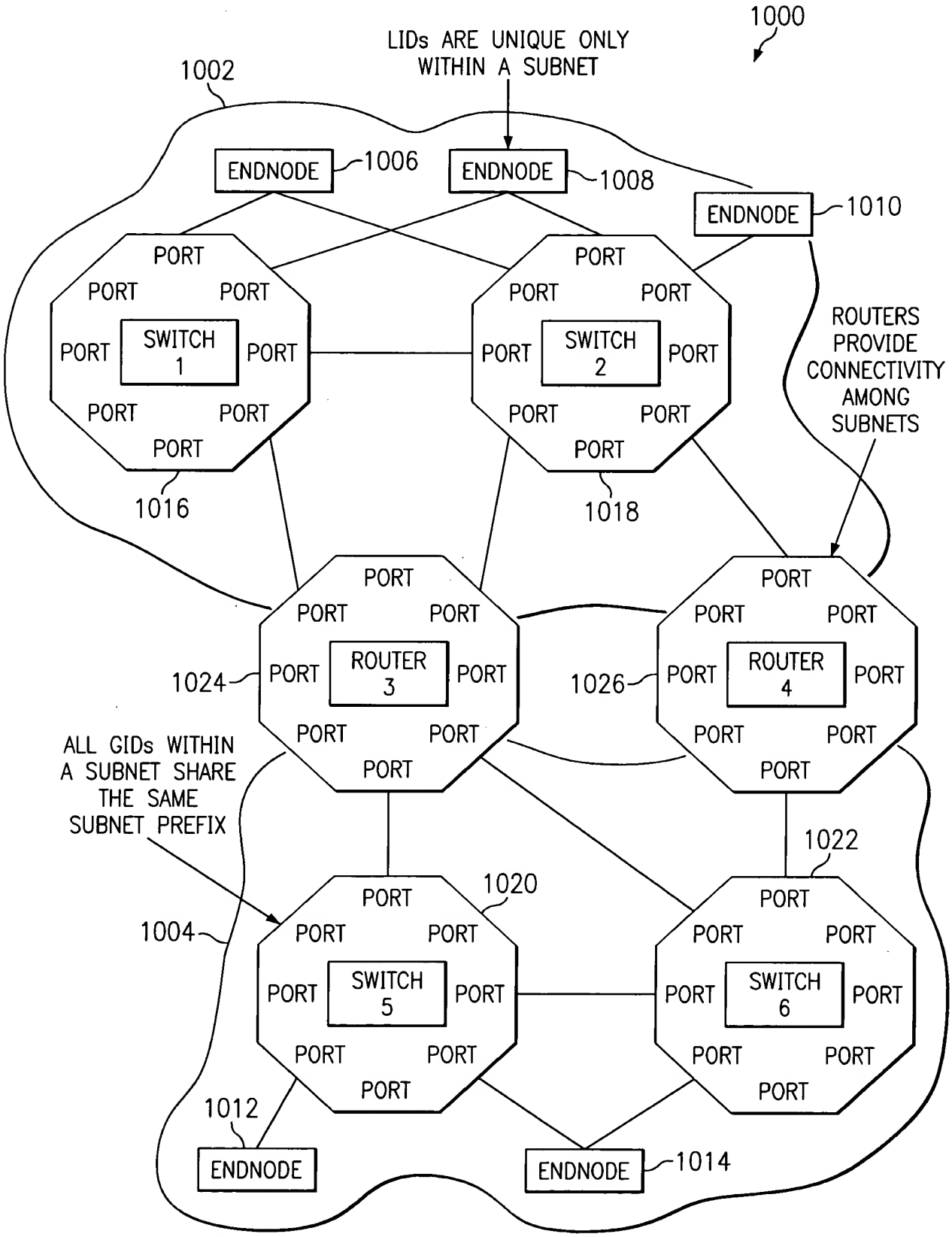

1000
LIDs ARE UNIQUE ONLY WITHIN A SUBNET
1002
ENDNODE
1006
ENDNODE
1008
ENDNODE
1010
PORT
SWITCH 1
SWITCH 2
ROUTERS PROVIDE CONNECTIVITY AMONG SUBNETS
1016
1018
ROUTER 3
ROUTER 4
1024
1026
ALL GIDs WITHIN A SUBNET SHARE THE SAME SUBNET PREFIX
1004
1020
1022
SWITCH 5
SWITCH 6
1012
1014

FIG. 10

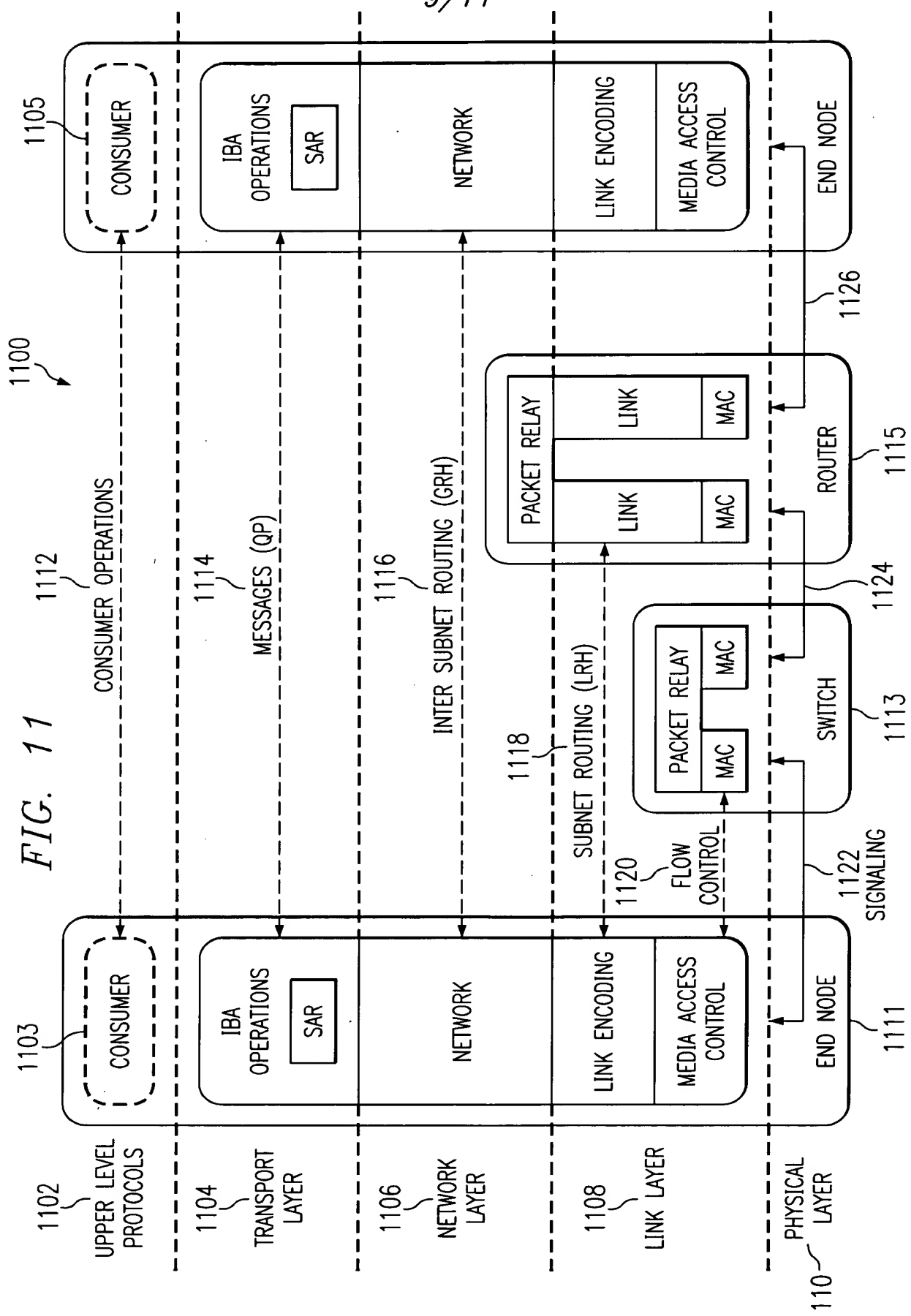

1100
1102
UPPER LEVEL PROTOCOLS
1104
TRANSPORT LAYER
1106
NETWORK LAYER
1108
LINK LAYER
1110
PHYSICAL LAYER
1103
CONSUMER
IBA OPERATIONS
SAR
NETWORK
LINK ENCODING
MEDIA ACCESS CONTROL
END NODE
1111
1105
CONSUMER
IBA OPERATIONS
SAR
NETWORK
LINK ENCODING
MEDIA ACCESS CONTROL
END NODE
1126
1112
CONSUMER OPERATIONS
1114
MESSAGES (QP)
1116
INTER SUBNET ROUTING (GRH)
1118
SUBNET ROUTING (LRH)
1120
FLOW CONTROL
1122
SIGNALING
PACKET RELAY
LINK
MAC
LINK
MAC
ROUTER
1115
1124
PACKET RELAY
MAC
MAC
SWITCH
1113

FIG. 11

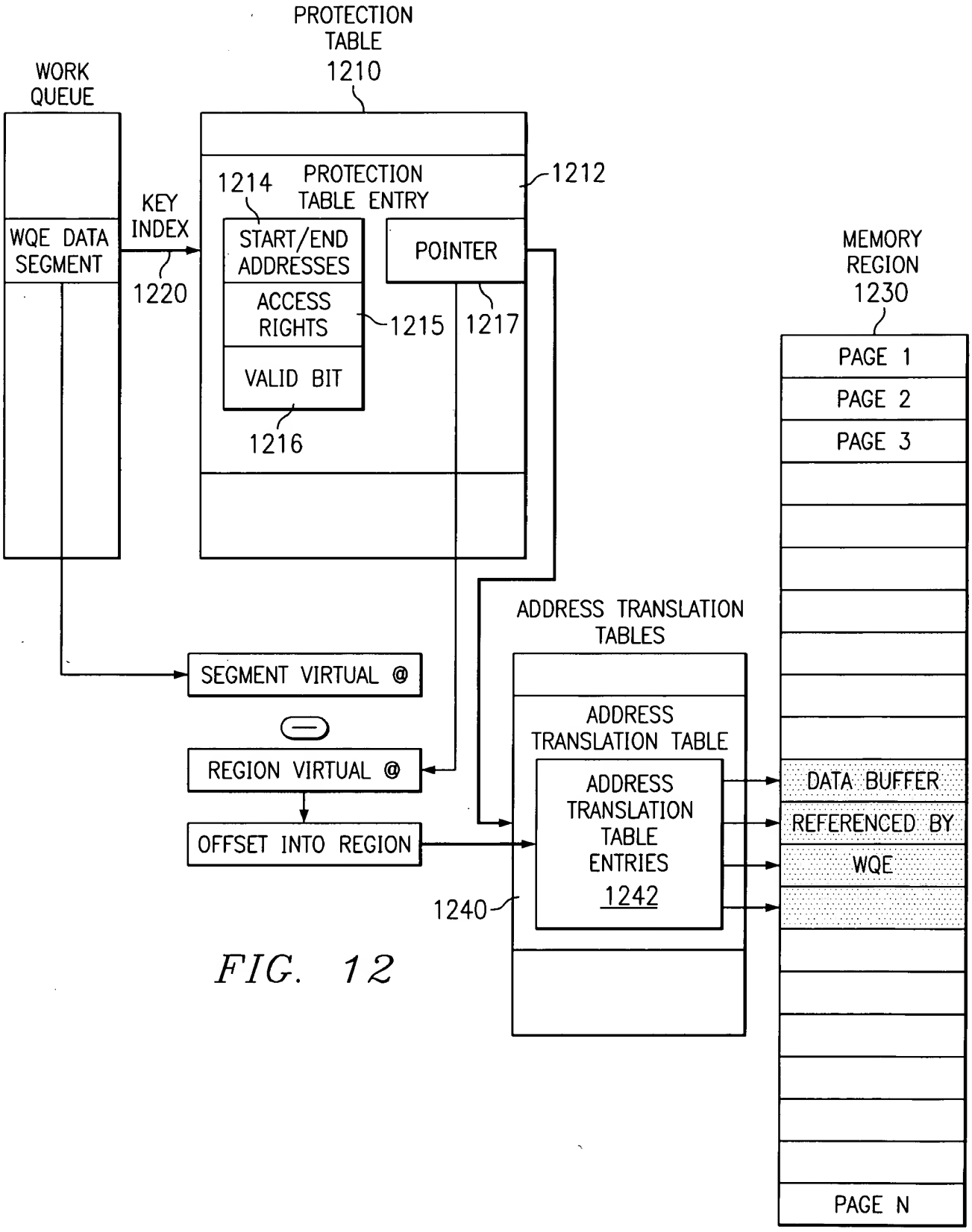
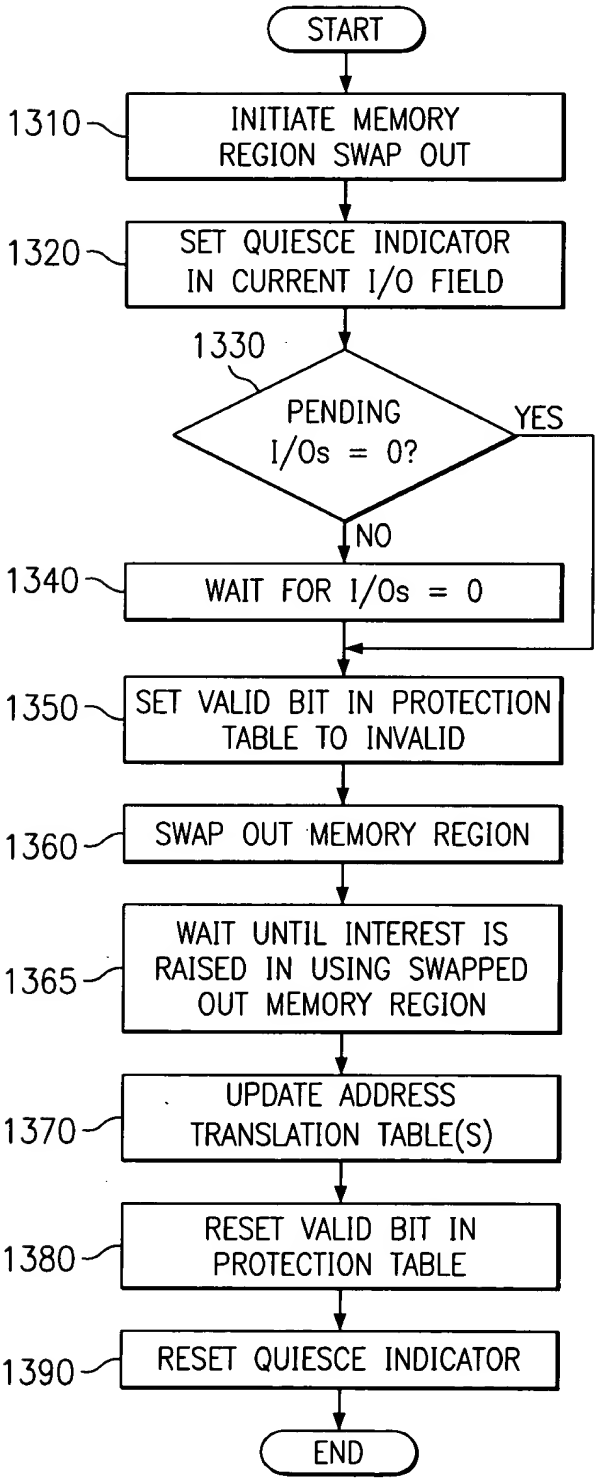

PROTECTION TABLE
1210
WORK QUEUE
WQE DATA SEGMENT
KEY INDEX
1220
1214
PROTECTION TABLE ENTRY
1212
START/END ADDRESSES
ACCESS RIGHTS
1215
VALID BIT
1216
POINTER
1217
MEMORY REGION
1230
PAGE 1
PAGE 2
PAGE 3
PAGE N
SEGMENT VIRTUAL @
REGION VIRTUAL @
OFFSET INTO REGION
ADDRESS TRANSLATION TABLES
ADDRESS TRANSLATION TABLE
ADDRESS TRANSLATION TABLE ENTRIES
1242
1240
DATA BUFFER
REFERENCED BY
WQE

FIG. 12

FIG. 13


START
1310
INITIATE MEMORY REGION SWAP OUT
1320
SET QUIESCE INDICATOR IN CURRENT I/O FIELD
1330
PENDING I/Os = 0?
YES
NO
1340
WAIT FOR I/Os = 0
1350
SET VALID BIT IN PROTECTION TABLE TO INVALID
1360
SWAP OUT MEMORY REGION
1365
WAIT UNTIL INTEREST IS RAISED IN USING SWAPPED OUT MEMORY REGION
1370
UPDATE ADDRESS TRANSLATION TABLE(S)
1380
RESET VALID BIT IN PROTECTION TABLE
1390
RESET QUIESCE INDICATOR
END